

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)特許出願公開番号

特開平5-257611

(43)公開日 平成5年(1993)10月8日

(51)Int.Cl.
G 0 6 F 3/06

識別記号 庁内整理番号
3 0 1 Z 7165-5B
K 7165-5B

F I

技術表示箇所

審査請求 未請求 請求項の数11(全 13 頁)

(21)出願番号 特願平4-356530

(22)出願日 平成4年(1992)12月22日

(31)優先権主張番号 8 1 2 3 3 9

(32)優先日 1991年12月23日

(33)優先権主張国 米国(US)

(71)出願人 592089054

エヌ・シー・アール・インターナショナル・インコーポレイテッド
アメリカ合衆国 45479 オハイオ、デイトン サウス バターソン プールバード 1700

(72)発明者 ティモシー ウェイン ジェントリー
アメリカ合衆国 67220 カンザス、ウィチタ、フラグスターフ 5917

(72)発明者 ジェラルド ジェイムズ フレディン
アメリカ合衆国 67212 カンザス、ウィチタ、ウエストポート 10203

(74)代理人 弁理士 西山 善章

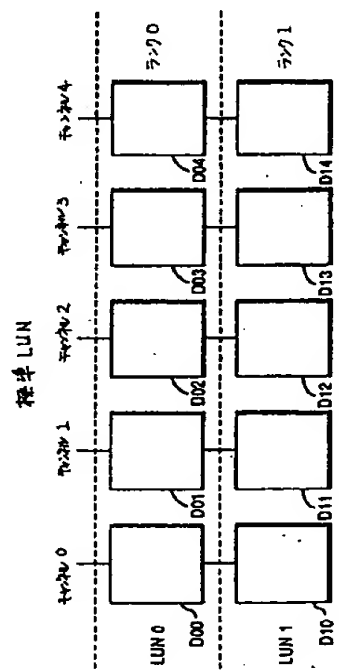
最終頁に続く

(54)【発明の名称】 ディスクアレーの区分け方法

(57)【要約】

【目的】 単一のディスクドライブセット上に複数の論理ユニットを支持することにより、または複数のディスクドライブセットを組合せて単一の論理ユニットを構成することにより、ディスクアレー論理ユニットの能力の向上を図る。

【構成】 ディスクアレーを物理保存ユニットとは別の論理保存ユニットに区分けする。1セットのドライブD00～D14が複数のパーティションに区分けされる。区分けされたドライブはグループ化され1つのディスクアレーとしてアドレス指定されかつ機能する論理ユニットを形成する。論理ユニット内のパーティションは論理アレー内のディスクドライブとしてアドレス指定されかつ機能する。これにより、1セットのドライブが複数の論理ユニットに分割され、その各々が独立したディスクアレーとして機能し、異なるRAIDレベルのスキームを用いてデータを保存する。



BEST AVAILABLE COPY

【特許請求の範囲】

【請求項1】 複数のディスクドライブにデータを保存する方法であって：上記複数のディスクドライブのうち少なくとも1つを複数のパーティションに区分けするステップと；上記パーティションおよび区分けされないディスクドライブのサブセットを1つの保存装置としてアドレス指定された論理ユニットに割り当てるステップとからなることを特徴とするディスクアレーの区分け方法。

【請求項2】 前記複数のディスクドライブのうち少なくとも2つを、それぞれ同じ態様で、複数のパーティションに区分けし、前記論理ユニットは前記ディスクドライブから区分けされた対応するパーティションからなることを特徴とする請求項1に記載のディスクアレーの区分け方法。

【請求項3】 前記パーティションおよび前記区分けされないディスクドライブが少なくとも2つのグループに分けられ、各グループが単一のディスクドライブとしてアドレス指定され；前記論理ユニットは少なくとも2つの前記グループからなり；前記論理ユニット内の各グループは、該論理ユニット内の単一のディスクドライブとして動作し；該論理ユニットはディスクアレーとして動作するように構成したことを特徴とする請求項1に記載のディスクアレーの区分け方法。

【請求項4】 複数のディスクドライブにデータを保存する方法であって：上記複数のディスクドライブの各々を、それぞれ同じ態様で、少なくとも2つのパーティションに区分けするステップと；上記パーティションを少なくとも2つの論理ユニットに割当て、各論理ユニットは上記複数のディスクドライブの各々から区分けした対応するパーティションを含むステップとからなり；上記論理ユニットの各々は独立したディスクアレーとして動作することを特徴とするディスクアレーの区分け方法。

【請求項5】 異なるRAIDレベルに従って、データを前記各論理ユニット内に保存することを特徴とする請求項4に記載のディスクアレーの区分け方法。

【請求項6】 複数のディスクドライブにデータを保存する方法であって：上記ディスクドライブを少なくとも2つのグループに割当てるステップと；選択されたグループを単一の論理ユニットに割当てるステップとからなり；上記論理ユニット内の各グループは該論理ユニット内で単一のディスクドライブとして動作し；上記論理ユニットはディスクアレーとして動作することを特徴とするディスクアレーの区分け方法。

【請求項7】 前記グループを少なくとも2つの論理ユニットに組織化し；異なるRAIDレベルスキーマに従って、データを各論理ユニット内に保存することを特徴とする請求項6に記載のディスクアレーの区分け方法。

【請求項8】 ディスクアレーを含むコンピュータシステムにおける上記ディスクアレー内にデータを保存する

方法であって：上記ディスクアレー内の個々のディスクを少なくとも2つのサブアレーに割当て、各サブアレーはデータを異なるRAIDレベルスキーマに従って保存するステップと；上記コンピュータシステムから受け取ったデータに応じて、上記サブアレーのうち1つを選択して該受け取ったデータを該サブアレー上に保存することを特徴とするディスクアレーの区分け方法。

【請求項9】 複数のチャンネルにわたって分配されかつ少なくとも1つのランクとして組織化された複数のディスクドライブからなるディスクアレー内で、該ランクは各チャンネルを1つのドライブに対応させたディスクドライブのセットからなる、ディスクアレー内のデータ保存方法において：上記ランク内の各ドライブを、それぞれ同じ態様で、少なくとも2つのパーティションに区分けするステップと；上記ランク内のディスクドライブから区分けされた対応するパーティションの各セットを対応する論理ユニットに割当てるステップと；各論理ユニットを異なるRAIDレベルスキーマに従って動作させデータを保存するステップとからなることを特徴とするディスクアレーの区分け方法。

【請求項10】 複数のチャンネルにわたって分配されかつ少なくとも2つのランクとして組織化された複数のディスクドライブからなるディスクアレー内で、該ランクの各々は各チャンネルを1つのドライブに対応させたディスクドライブのセットからなる、ディスクアレー内のデータ保存方法において：第1のランクに属する各ディスクドライブを第2のランクに属するディスクドライブに連携させるステップと；上記論理ユニット内の対応するディスクドライブの各ペアを単一のディスクドライブとしてアドレス指定するステップと；選択されたRAIDレベルスキーマに従って上記論理ユニットを動作させデータを保存するステップとからなることを特徴とするディスクアレーの区分け方法。

【請求項11】 ディスクアレーにデータを保存する方法であって：上記ディスクアレーを複数の論理ユニットに区分けするステップと；異なるRAIDレベルスキーマに従って各論理ユニット内にデータを保存するステップとからなることを特徴とするディスクアレーの区分け方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、コンピュータシステムのためのディスク保管装置に関し、特にRAID (Redundant Arrays of Inexpensive Disks)等のディスクアレー保管装置に関するものである。

【0002】

【従来の技術】 RAID保管装置は、今までの大型で高価なディスクドライブに代えてこれからのコンピュータシステムにおいて用いられるために開発された。RAID保管装置は、現在パーソナルコンピュータやワークス

ーションで用いられている5.25インチや3.5インチのディスクドライブ等のような小型で安価なハードディスクドライブのアレーを備えている。ディスクアレー装置は数年前より存在するが、近年その小型ディスクドライブの信頼性や操作性の向上およびコストの低下に伴いRAID装置への関心が高まっている。

【0003】複数のディスクドライブを組合せて多種類のディスクアレーを構成することができる。各々異なる作用効果および欠点をもったいくつかのディスクアレーが、「低価格ディスクの重複アレー(RAID)用ケース」、デビット エー バターソン、ガース ギブソン、およびランディ エイチ キャッツ著；カリフォルニア大学リポートNo. UCB/CSD 87/391, 1987年12月に開示されている。従来技術として参照したこの文書には、単一で大型の磁気ディスクに比べた場合のディスクアレーの作動、信頼性、消費電力およびサイズ等についての改良点が開示されている。

【0004】この文書に記載されたRAIDレベル1は、データ保存用のN枚のディスクとこのデータ保存ディスクの情報をコピーして保存するN枚のミラーディスクを具備する。このRAIDレベル1の書き込み動作はデータを2枚のディスクに書き込むことが必要であり、最初のディスクに書き込んだデータが2番目のミラーディスクに書き込まれる。データを読み出す場合には、いずれのディスクからでも読み出すことができる。

【0005】図1は4つのドライブを含むRAIDレベル1システムを示す。4つのドライブはデータ1、ミラー1、データ2、ミラー2と表示されている。各ディスクドライブの下に示すブロックはディスク上でのデータ保存形式を示す。

【0006】RAIDレベル1は、高レベルの重複性と高いデータ交換性を有し、書き込みペナルティは低減され回復ペナルティはない。このRAIDレベル1ではデータの可用性は非常に大きい。重複ドライブに起因するコストと利用可能なデータ保存能力の点で問題があり、この点は、RAIDレベル3およびレベル5により改良されている。

【0007】RAIDレベル3システムは、N+1枚のディスクからなる1又はそれ以上のグループにより構成される。各グループ内でN枚のディスクはデータ保存用に用いられ、他のディスクはパリティ情報を保存するために用いられる。このRAIDレベル3の書き込み動作の間、各データのブロックはN枚のデータディスクに保存されるためにN個に分割される。対応するパリティ情報は専用のパリティディスクに保存される。

【0008】データが読み出される場合には、N枚のディスク全てにアクセス可能でなければならない。パリティディスクはディスク障害が発生したときに情報を再生するために用いられる。

【0009】図2は5つのドライブを備えたRAIDレ

ベル3システムを示す。5つのディスクドライブはそれぞれデータ1～データ5と記されている。データはデータ1～データ4のディスクにストライプされ、各データディスクは保存されるデータの一部を保持する。データ1～データ4のディスクドライブに保存されたデータのビット側排他的ORを介して発生したパリティ情報はデータ5のドライブに保存される。

【0010】RAIDレベル3はバイトレベル又はワードレベルでストライプするデータを発生し、データ転送速度は非常に速くまた書き込み動作あるいはデータ回復動作のペナルティはない。RAIDレベル3システムは、例えば画像、モデリング、およびシミュレーション決定支援、強調グラフィックおよび画像処理、科学コンピュータ処理およびCAD/CAMへの適用等の大規模なファイル転送に対して用いられれば効果的である。

【0011】図3はRAIDレベル5システムを示す。このレベル5システムもN+1枚のディスクからなる1又はそれ以上のグループにより構成される。RAIDレベル5がレベル3と異なる点は、レベル5システムにおいては、保存すべきデータが複数のディスクに保存されるために1つ又はそれ以上のデータブロックからなるさらに多くの部分に分割されていることである。さらに、各グループはN+1枚のディスクを含んでいるが、各ディスクはデータ保存ブロックとパリティ情報保存ブロックを有している。即ち、パリティ情報は、特定のパリティディスクドライブに保存されるのではなく、N+1のドライブに分配される。

【0012】RAIDレベル5の書き込みには少なくとも2枚のディスクへのアクセスを必要とする。しかしながら、グループへの各書き込み動作においては、同じ専用のパリティディスクへのアクセスを必要としない。各ディスクはパリティ情報を保存するためのいくつかのブロックを有している。同様に、読み出し動作においても、N枚のデータディスクのうち1枚のディスクへのアクセスしか必要としない。RAIDレベル3システムではパリティデータはディスク障害が発生したときに情報を再生するために用いられる。

【0013】RAIDレベル5は、システムブロックサイズでデータのストライプを行い、全ドライブに多くパリティを分配し、データ交換作用が改善されているが、上位の書き込みペナルティがある。このRAIDレベル5システムは、高いI/O速度と小さいブロックサイズが要求されるスーパーコンピュータあるいはトランザクション処理に対して適用すれば最も効果的である。RAIDレベル5システムは、航空機や自動車の予約センターでのオンライン処理や自動預金支払機あるいはPOSシステムおよびデータベース等に通用することが最も理想的である。

【0014】図4は、別のディスクアレー構成であるRAIDレベル0を示す。このアレーはデータ保存用のN

10

20

30

40

50

5

枚のデータディスクを備える。アレーコントローラは各ドライブに各別々にアクセスし、5つの異なる位置でNの読み出し又は書き込み動作を同時に行う。これにより透過ロードバランスが達成され1枚のディスクドライブの動作能力が向上する。RAIDレベル0にはパリティの発生や保存はなく、従ってRAIDレベル1、3、5のようにデータの回復あるいは再生作用はない。

【0015】RAIDレベル0は、システムブロックサイズでデータをストライプし、処理能力が向上し、トランザクション処理性能が高まり、書き込みペナルティはない。しかしながらデータの回復や拡張データの利用ができない。このRAIDレベル0は追加性能を要する処理に対し適用すれば効果的であるが、他のRAIDレベルで利用できるデータが適用できない。

【0016】前述のように、異なるRAIDレベルシステムに対し異なるコンピュータを適用すれば効率的である。例えば、RAIDレベル1システムは高いデータ可用性が要求される処理に適用することが最適であり、RAIDレベル3は大規模なファイル転送が要求される処理に適用することが最適であり、RAIDレベル5は高いI/O速度と小さいブロックサイズが要求される処理に適用することが適当であり、RAIDレベル0は追加性能が要求される処理に適用することが最適であるが他のレベルで適用されるデータが利用できない。従って、複数の処理に対し適用できる処理装置に対しては、同時に2以上のRAIDを利用できるアレーが従来の単一RAIDレベルのアレーに比べ効果が大きい。

【0017】

【発明が解決しようとする課題】小型のディスクドライブの能力が向上し、オペレーティングシステムや適用ディスクサイズの制限にかかわらずディスクアレーの能力が増加したため、ディスクアレーを物理的に2又はそれ以上の小さい論理アレーに区分する方法が要求されている。

【0018】従って、本発明の目的は、新規で改良されたディスクアレーの区分け方法を提供することである。本発明の別の目的は、複数のRAIDレベルを同時に支持できるディスクアレーを構成する方法を提供することである。本発明のさらに別の目的は、単一のディスクドライブセット上に複数の論理ユニットを支持する方法を提供することである。

【0019】本発明のさらに別の目的は、複数のディスクドライブセットを組合せて単一の論理ユニットを構成する方法を提供することである。本発明のさらに別の目的は、ディスクアレー論理ユニットの能力の向上を図った新規なアレー仕切方法を提供することである。

【0020】

【課題を解決するための手段】前記目的を達成するため、本発明では、ディスクアレーを区分してアレー内の物理的保存ユニットとは別に論理保存ユニットを形成す

6

る。本発明の一形態によれば、アレー内の個々のドライブは多数のパーティションに区分けされる。2又はそれ以上のドライブからのパーティションはグループ化されて単一の論理ユニットを構成する。この論理ユニットは独立したディスクアレーとしてアドレス指定されかつ作動する。論理ユニット内のパーティションは論理アレー内のディスクドライブとしてアドレス指定されかつ作動する。従って、ディスクドライブの単一セット又はランクは、2又はそれ以上の論理ユニットに分割され、その各々は1つの独立したディスクアレーとして作用しかつ各々別のRAIDレベルを用いてデータを保存する。

【0021】本発明の別の実施例においては、アレーのディスクドライブは2又はそれ以上のグループに分けられ、各グループは2又はそれ以上のドライブを含んでいる。これによりアレーは単一のディスクドライブのアレー能力の2倍又はそれ以上の能力を持つように形成される。

【0022】上記2つの実施例を組合せて複数のディスクドライブからいくつかの論理アレーを形成してもよい。この論理アレーのうちいくつかはディスクドライブのパーティションから形成され、いくつかは区画されないディスクドライブのグループを組合せて形成され、その他のアレーは区画されないドライブとドライブのパーティションとを組合せて形成される。各論理アレーは異なるスキーマを用いてデータを保存してもよい。

【0023】

【実施例】図5は、10枚のディスクドライブを含むディスクアレーを示し、これらのディスクドライブはD00～D04およびD10～D14で示されている。10のドライブは物理的に2つのグループ（ランク）に配列される。この2つのグループRANK0およびRANK1は各々5つのドライブを含んでいる。2つのランクの対応するドライブ、ドライブD01とD11等、は同じバス（又はチャンネル）を介してアクセスされる。図5にはチャンネル0～4の5つのチャンネルが示されている。

【0024】単一の保存装置としてシステムに示されたアレーの一部を形成するディスクアレー論理ユニットは、多数のチャンネル又は単一のバスを介したディスクドライブのセットとして構成される。図5に示すアレーにおいては、RANK0およびRANK1の各々はLUNと表示した論理ユニットを構成する。各LUNのキャパシティは、パリティ情報の予約キャパシティのいずれよりも小さい個々のディスクドライブのキャパシティの合計である。本発明に係る方法においては、大型で高価なディスクをアレー内に組み込むことなくディスクアレーの論理ユニットのキャパシティを増加させる。

【0025】図6は、本発明方法に従って図5のアレーを配列しRANK0およびRANK1からなる単一の論理ユニットを構成したアレーを示す。スーパーLUNと

表示したこの新しい論理ユニットは図5のLUN0又はLUN1のいずれのキャパシティの2倍のキャパシティを有する。このスーパーLUNは、前述のいかなるRAIDレベルでもサポートすることができ、この場合ディスクドライブは同じチャンネル上にありサポートされたRAIDレベル内で単一のドライブとして作用する。同じチャンネル上にあるディスクドライブはコントローラにより多数のドライブとして管理されすべてのドライブに対するロードバランスが得られる。

【0026】図7は、本発明方法に従って、LUN0～LUN4の4つの論理ユニットを形成するように配列された図5のアレーを示す。各論理ユニットはランク内の各ディスクドライブの対応部分にコマンドを入力する。サブLUNと表示したこの新規で小型の論理ユニットは、大容量論理ユニットとの連携不可能なオペレーティングシステムと協働してアレーコントローラを作動させることができる。各サブLUNは前述のRAIDレベルのどれでもサポート可能である。このようにして、単一のドライブランクを2又はそれ以上の論理ユニットに分割し、2又はそれ以上のRAIDレベルをサポートすることができる。

【0027】図8は、図5のアレーの別の配列構成例を示す。この例では各ディスクドライブが論理ユニットとして作用する。LUN0～LUN9までの10の論理ユニットが開示してある。図5から図8までのLUN形態

は図の例に限定されず、例えば、図示した2つのディスクドライブのランクから、ドライブD00～D04のアナログ部から形成したサブLUNとRANK0の残りのキャパシティとRANK1とを組合せて形成したスーパーLUNの両方を構成してもよい。

【0028】ディスクアレーの論理配列は、ディスク配列エリアに格納されたいつつかのテーブルに含まれる。これらのエリアは、ホストシステムおよびアレーコントローラに利用されシステムのユーザに特定された方法でデータを記憶したり引出したりする。

【0029】各論理ユニットの構造は論理ユニット情報ブロック内に記述される。このブロックは、LUNで用いるRAIDレベルのスキーマや、LUNのブロックサイズや、LUNのセグメントサイズや、後述の各種パラメータの情報を格納している。この論理ユニット情報ブロックは、さらに1又はそれ以上のランク情報ブロック内にポインタを含んでいる。このランク情報ブロックは物理的なディスクドライブランクの構造を記述する。ディスク情報ブロックのセットがすべてのチャンネル上のすべてのディスクドライブの構造を記述する。論理ユニット情報ブロックのサンプルを下記の表1に示す。また、ランク情報ブロックおよびディスク情報ブロックの例を下記の表2および表3にそれぞれ示す。

【0030】

【表1】

バイト	フィールド
0	LUN番号
1	LUNタイプ
2	RAIDレベル
3	ランク数
4-5	フラグ
6	LUN状態
7	セグメント0サイズ
8-11	LUNセグメントサイズ
12-15	最終LUNデータブロック
16-19	LUNブロックサイズ
20-23	スタートディスクブロック
24-27	ディスク容量
28-31	ランク情報
32-35	RAID
36-39	ホスト情報
40-43	予約
44-47	ページ情報
48-49	再生量
50	再生回数
51	予約

【0031】論理ユニット情報ブロックにおいて、LUN番号のフィールドはLUNを特定する。LUNタイプのフィールドはLUNの型式を特定する。例えば、このフィールドが0の値であれば通常のLUNを示し、ドライブ上の利用可能な全スペースを占有する。このフィールドが1の値であればサブLUNを示しドライブ上の1部のスペースを占有する。また2の値であればスーパーLUNを示し複数のランクにまたがる。RAIDレベルのフィールドは、LUNで用いるRAIDレベルの型式が0、1、3、または5のいずれかを示す。ランク数のフィールドはLUNで用いるランクの数を示す。LUNフラグのフィールドは、LUNに対するコマンドとオペレーションのオプションを定義するために用いられる。

LUN状態のフィールドは、LUNの現在の状態を示すために用いられる。LUNは1日の処理作業の開始状態が整わないとアクセスできないようにしてもよい。セグメント0サイズのフィールドは、セグメント0に割り当てられたブロックの数を示す。このフィールドが含まれる理由は、LUN上の最初の数個のブロックにアクセスする最適セグメントサイズがこのLUNの他のすべてのブロックにアクセスする最適のセグメントサイズと異なる場合があるからである。

【0032】LUNセグメントサイズは、単一のディスク上でグループ化された論理ブロックの数である。最終LUNデータブロックおよびLUNブロックサイズのフィールドは、ホストへの転送のためのデータ容量および

論理ブロックサイズを特定する。LUNのスタートディスクブロックおよびLUNの各ドライブに含まれるディスクブロックの数は、スタートディスクブロックおよびディスク容量のフィールドにそれぞれ示される。ランク情報のフィールドは、LUNに含まれるすべてのランク*

*のためのランク情報ブロックに対するポインタを含んでいる。この後にサンプルランク情報ブロックが続く。

【0033】

【表2】

ランク情報ブロック

バイト	フィールド
0	ランク番号
1	ディスク/ランク
2	状態
3	不良ディスク
4 - 23	ディスク情報
24 - 27	ランク内論理ブロック
28 - 31	スタート論理ブロック
32 - 35	エンド論理ブロック
36 - 39	ランク内セグメント
40 - 43	スタートセグメント
44 - 47	エンドセグメント
48 - 51	ランク内パリティグループ
52 - 55	スタートグループ
56 - 59	エンドグループ
60 - 63	スタートディスクセクター

【0034】ランク情報ブロックは、ランクの構造を示す。このランク情報ブロック内で、ディスク/ランクのフィールドはランク内に含まれるディスクの数を示す。ディスク情報のフィールドは、LUNで用いる個々のディスクドライブを記述したディスク情報ブロックのセットに対するポインタを含んでいる。ランク容量、ランクのスタートおよびエンドセグメント、ランク内のパリティグループ数、スタートおよびエンドパリティグループ、およびランクスタートディスクセクターに関する情報はランク内のバイトオフセット24～63に格納される。この後にサンプルディスク情報ブロックが続く。

【0035】

【表3】

13
ディスク情報ブロック

バイト	フィールド
0	チャンネル
1	ディスク
2	ディスク LUN
3	日程スタート状態
4	ドライブ状態
5	ファームウェアバージョン
6	ディスクオブションフラグ
7	予約
8 - 9	ディスクセクターサイズ
10 - 11	セクター/トラック
12 - 13	ECCセクターサイズ
14 - 17	スタートデータセクター
18 - 21	エンドデータセクター
22 - 25	スタート保持セクター
26 - 29	エンド保持セクター
30 - 33	エンドディスクセクター
34 - 37	連携 LUN
38 - 41	ドライブ対象

物理アレーページ

バイト	モード選択レイアウト	モード感知レイアウト
0	ページコード 2 A	ページコード 2 A
1	ページ長 (240 dec)	ページ長 (240 dec)
2 - 241	ドライブ状態	ドライブ状態

【0039】物理アレーページは、ディスクアレーコントローラに接続された使用可能性のあるチャンネルおよびID位置のための1バイト状態を含む。各バイトは、アレー内での一意ドライブ用の特別状態の値を含んでいる。各ドライブの状態情報は、ドライブが正しく動作し

(8)

特開平5-257611

14

*【0036】アレー内の各ドライブは、上記表のブロックに示すように、ディスク情報ブロックにより1つ1つ特定される。

【0037】ディスクアレーは、アレーの物理面および論理面を定義する2つのモード選択ページおよびテーブルを用いて形成される。前述のディスク、リンクおよびLUNの情報ブロックに含まれるフィールドは、これら2つのモード選択ページから引出される。物理アレーページと論理アレーページを示すこれらの2つのページの各々はアレー動作を成功させるために必要なキーパラメータを含んでいる。物理アレーページは、アレーの全ての物理面をその論理組織形態にかかわらず制御する。以下の表4は、最大240のドライブ位置を有するアレー用の物理アレーページの組織を示す。

【0038】

【表4】

20

30

*

ていれば最適状態を示し、ドライブが交換されてアレーが再構成されているときには再構成状態が示される。モード選択オペレーション中に、ホストはアレーの物理面に関する情報を追加、消去又は修正することができる。このモード選択オペレーション中に現在の情報がレポー

50

トされる。

【0040】論理アレーページはディスクアレーの論理組織を制御する。このパラメータは、論理装置（ユニット）のRAIDレベル、論理装置のセグメントサイズ、* 論理アレーページ

* 再構成の回数、および再構成の量のデータを含む。以下の表5は論理アレーページの例を示す。

【0041】

【表5】

バイト	モード選択レイアウト	モード感知レイアウト
0	ページコード2B	ページコード2B
1	ページ長(130dec)	ページ長(130dec)
2	動作	LUN状態
3	RAIDレベル	RAIDレベル
4	LUNタイプ	LUNタイプ
5-8	LUNブロックサイズ	LUNブロックサイズ
9-10	予約	ドライブセクターサイズ
11-14	予約	ブロックLUN番号
15-18	予約	予約
19-22	LUNセグメントサイズ	LUNセグメントサイズ
23	LUNセグメント0サイズ	LUNセグメント0サイズ
24-25	LUNフラグ	LUNフラグ
26-29	予約	再構成ブロック完了
30	再構成回数	再構成回数
31-32	再構成量	再構成量
33-36	予約	予約
36-67	ディスクビットマップ	ディスクビットマップ
68-131	構成テーブル	構成テーブル

【0042】特定の論理装置のRAIDレベルはバイトオフセット3により制御される。LUNのタイプ、即ち図5に示す通常のLUNか、図6に示すスーパーLUNか又は図7に示すサブLUNのいずれであるかはバイトオフセット4により特定される。LUNブロックサイズ、LUNセグメントサイズおよびLUNセグメント0サイズについては前述の表1の場合と同じである。バイトオフセット30および31～32でそれぞれ特定する再構成回数および再構成量は、単一のドライブに障害が発生した場合にアレーが再構成される速さを示す。バイトオフセット36～67のディスクビットマップはLUNで用いられる物理ドライブを特定する。バイトオフセ

ット68～131の構成テーブルは、ドライブをストライプするデータの順序、あるいはRAID1の場合には、どのドライブがミラードライブになるかを決定する。

【0043】モード選択オペレーションでは、ホストはLUNを追加、消去あるいは修正することができる。現在のLUNの定義はモード感知オペレーションを介してレポートされる。

【0044】情報ブロックおよびアレーテーブルを用いてユーザの定義する論理装置が構成されることは、前記詳細な説明から当業者に容易に理解できる。本発明により、新規に改良されたディスクアレーの区分け方法が提

供される。この方法においては、複数のRAIDレベルがディスクアレーにより同時にサポートされ、単一のディスクドライブセットが複数の論理ユニットに区分けされ、あるいは複数のドライブセットが組合されて単一のドライブセットとして作用する単一の論理ユニットを構成する。

【0045】なお、本発明は前述の実施例に限定されず各種変更態様が可能である。

【0046】

【発明の効果】以上説明したように本発明においては、ディスクアレーを論理保存ユニットと物理保存ユニットに分けて区分けし、論理ユニットを独立したディスクアレーとして作用させているため、小型ディスクドライブの容量を増加させることができる。

【図面の簡単な説明】

【図1】 2つのデータおよび2つのミラーディスクドライブを含むRAIDレベル1のアレーを示すブロック図である。

【図2】 4つのデータディスクドライブと1つのパリティディスクドライブを含むRAIDレベル3のアレー*20

*を示すブロック図である。

【図3】 5つのディスクドライブを含むRAIDレベル5のアレーを示すブロック図である。

【図4】 5つのディスクドライブを含むRAIDレベル0のアレーを示すブロック図である。

【図5】 各々5つのディスクドライブからなる2つのランクに分けられた10のドライブのアレーを示す構成説明図である。

【図6】 本発明の一実施例に係る単一の論理ユニットとして組立られた2つのランクからなる10のドライブのアレーを示す構成説明図である。

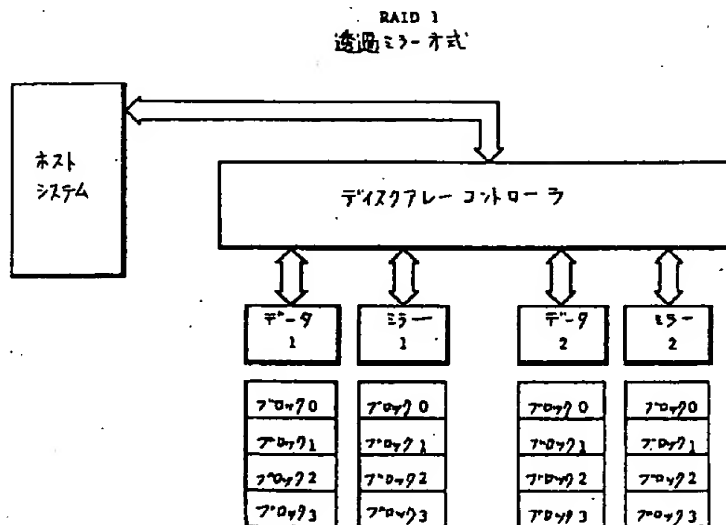
【図7】 各々のランクを本発明に係る2つの論理ユニットに分割した2つのランクからなる10のドライブのアレーを示す構成説明図である。

【図8】 各ディスクドライブが本発明に係る論理ユニットを構成する2つのランクからなる10のドライブのアレーを示す構成説明図である。

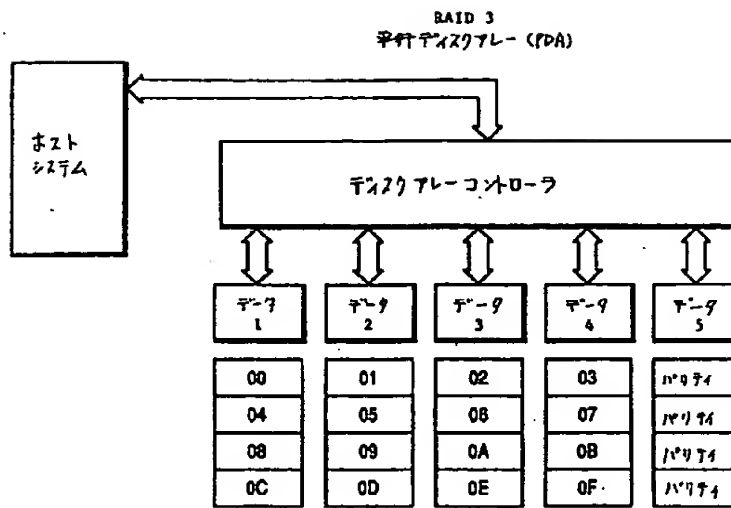
【符号の説明】

D00～D04、D10～D14；ディスクドライブ。

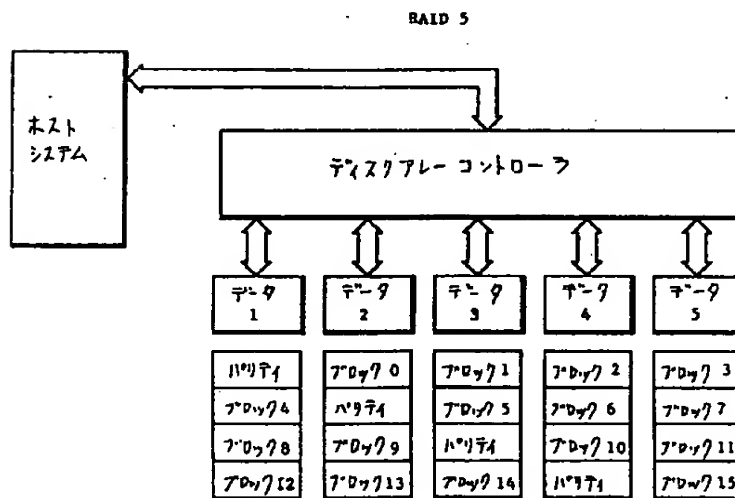
【図1】



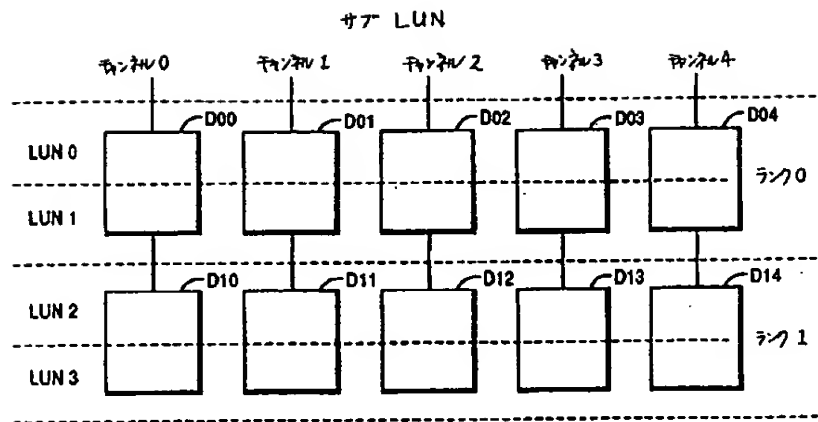
【図2】



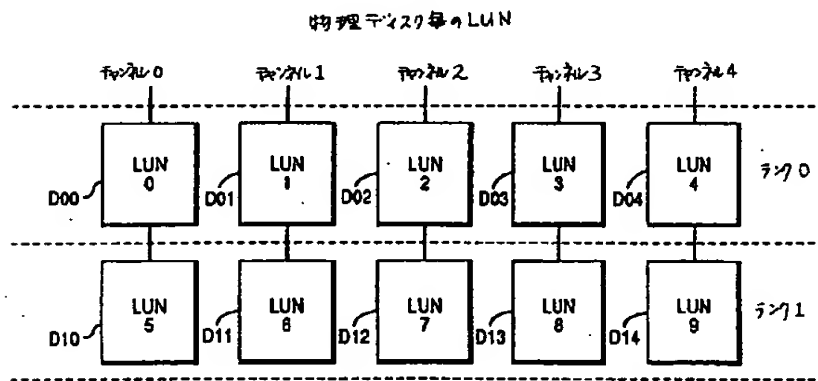
【図3】



【図7】



【図8】



フロントページの続き

(72)発明者 ダニエル アンソニー リードゥル
 アメリカ合衆国 67002 カンザス、アン
 ドヴァー、ダブリュー、ダグラス 425

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.